

Statistical Modeling for Spatio-Temporal Degradation Data

Xiao Liu, Kyongmin Yeo, Jayant Kalagnanam

IBM T.J. Watson Research Center

Abstract

This paper investigates the modeling of an important class of degradation data, which are collected from a spatial domain over time; for example, the surface quality degradation. Like many existing time-dependent stochastic degradation models, a special random field is constructed for modeling the spatio-temporal degradation process. In particular, we express the degradation at any spatial location and time as an additive superposition of two stochastic components: a dynamic spatial degradation generation process, and a spatio-temporal degradation propagation process. Some unique challenges are addressed, including the spatial heterogeneity of the degradation process, the spatial propagation of degradation to neighboring areas, the anisotropic and space-time non-separable covariance structure often associated with a complex spatio-temporal degradation process, and the computational issue related to parameter estimation. When the spatial dependence is ignored, we show that the proposed spatio-temporal degradation model incorporates some existing pure time-dependent degradation processes as its special cases. We also show the connection, under special conditions, between the proposed model and general physical degradation processes which are often defined by stochastic partial differential equations. A numerical example is presented to illustrate the modeling approach and model validation.

Key words: *Degradation, Stochastic Process, Convolution, Spatio-Temporal Statistics, Random Field.*

1 Introduction

Over the past two decades, the statistical modeling of degradation data has become an important topic in reliability engineering and industrial statistics. As lifetime data are typically censored for highly reliable product, degradation data provide a richer source of information for reliability analysis. The advantages of using degradation data for reliability assessment has been thoroughly discussed (Lu, Meeker and Escobar, 1996; Meeker, Escobar and Lu, 1998), and various successful applications are available in the literature (Bae and Kvam, 2004; Peng and Tseng, 2010; Chen and Tsui, 2012; Doganaksoy and Hall, 2013; Li and Meeker, 2013; Ye and Chen, 2014; Bian et al., 2015; Hong et al., 2015).

1.1 Degradation Models based on Stochastic Processes

In the literature, most previous work has been concerned with the modeling of pure time-dependent degradation processes; for example, the physical degradation process of a device over time, or, the overall performance degradation of a complex system. Meeker and Escobar (1998) described the General Path Models (GPM) with the unit-to-unit variation being captured by a random-coefficients model. Following the idea of GPM, Bae and Kvam (2004) proposed a nonlinear degradation model for vacuum fluorescent displays, Liu and Tang (2010) developed a nonlinear hot-carrier induced degradation model for MOS Field Effect Transistors, and Hong et al. (2015) proposed a statistical method for degradation data modeling with dynamic covariates and presented an application to outdoor weathering data. It is worth noting that, the modeling of degradation data under dynamic environments has also received much attention in recently years (Liao and Tian, 2012; Zhou, Serban and Gebraeel, 2014; Bian et al., 2015; Hong et al., 2015).

There exists another important class of degradation models, which rely on stochastic processes as approximations to the real-world degradation. In other words, the degradation process is modeled by a measurable real-valued function,

$$Y : [0, \infty) \rightarrow \mathbb{R}, \quad (1)$$

known as a stochastic process with the quantity $Y(t)$ representing the value of degradation at time t . Such an idea can be traced back to the early work of Bhattacharyya and Fries (1982) and Doksum and Hoyland (1992). Commonly used stochastic processes include the Wiener process (Tseng and Peng, 2004), Gamma process (Singpurwalla, 1995; Lawless and Crowder, 2004), Inverse Gaussian process (Ye and Chen, 2014), which are popular members of the family of Lévy processes. The fundamental idea of using Lévy processes, with independent and stationary increments, is rooted in the assumption of an additive accumulation of degradation, i.e., every degradation increment can be seen as an additive superposition of a number of *stationary and independent* small increments. Once the probability distribution for each small degradation increment is specified, the stochastic degradation process is uniquely determined. The use of stochastic processes allows us not only to model the temporal correlation structure of a degradation process, but also to leverage the mathematical properties, such as the sample path properties and transition density, which have been well established for the stochastic process adopted. For example, when the Wiener process is used, it is well-known that the time for the degradation to hit a given level for the first time is inverse Gaussian distributed, making it extremely convenient to model the time to some pre-defined events such as failure or preventive maintenance.

Another advantage to rely on stochastic processes is due to their strong connection to stochastic Partial Differential Equations (PDE), through which many physical degradation models are defined. For example, the Paris-Erdogan law that describes the fatigue crack growth, the diffusion of lithium ions in solid and electrolyte phases that causes the output degradation of lithium-ion batteries, the gradual decomposition of polymer microspheres for pharmaceutical drug delivery over extended periods of time, etc. In fact, since existing stochastic degradation models are based on the basic idea that every degradation increment is an additive superposition of a number of stationary and independent small increments, most of the existing stochastic degradation models can be written in the form of a differential equation with the general expression, $dY(t) = udt + dB(t)$, where $dY(t)$ is the degradation increment over an infinitesimal time interval dt , u is the instantaneous degradation rate, and $B(t)$ is some stochastic process. It is only the choice of $B(t)$ that differentiates a stochastic

degradation model from another. One may refer to Nikulin et al. (2010) and Ye and Xie (2014) for a comprehensive review of stochastic degradation models under this framework. We will revisit the representation of a stochastic degradation process by PDE in Sections 3 and 4.

1.2 Degradation in Space and Time

In this paper, we investigate the modeling of an important class of degradation data which are collected not only over time but also from a spatial domain. In particular, we extend the existing time-dependent degradation process, $Y : [0, \infty) \rightarrow \mathbb{R}$, to a space-time process,

$$Y : \mathbb{R}^d \times [0, \infty) \rightarrow \mathbb{R}, \quad (2)$$

where Y is now a spatio-temporal random field, and $d \in \mathbb{N}^+$ is the dimension of the spatial domain.

Degradation data of this kind usually exhibit complex correlation structure in space and time, making it no longer appropriate to model the data collected from different spatial locations independently using existing methods reviewed in Section 1.1. Some examples of spatio-temporal degradation include the aging of asphalt pavement on highways, the corrosion of oil and gas pipelines, the bulk erosion when water diffuses into a polymer structure, the surface quality degradation/drift of a manufacturing process such as the Wafer Intelligence Scanner (WIS) data that measure the RGB color of a silicon wafer layer. In Appendix B, we also show a real spatio-temporal model that describes the degradation of polymer microspheres composed of poly(D,L-lactic-co-glycolic acid) (Versypt et al., 2015).

Figure 1, as a motivating example, shows the degradation measured on a two-dimensional surface over 9 equally spaced time intervals (the darker the color, the higher the degradation). Due to the confidentiality agreement, the actual application is not mentioned throughout this paper and the degradation data are multiplied by a common scaling factor. In this example, the degradation data are measured by a surface scanner and aggregated to a 21×21 pixel array, and observations are available from time 1 to time 20. At time 1 (the initial condition), the measured degradation over the entire surface is relatively low, and the spatial variation across different pixels appears to be random with no special spatial pattern. Starting from time 3 and time 5, three regions, aligned horizontally in the center of the surface, are observed with higher degradation level. Both the degradation level and the size of these three regions gradually grow over time. And also, we note that the degradation propagates mainly following the south-to-north direction.

To our knowledge, the modeling of degradation data in both space and time has not been well studied in the literature. Compared to the modeling of pure time-dependent degradation data, the modeling of spatio-temporal degradation data of this kind poses some unique challenges, which are summarized as follows:

- **spatial propagation.** As already noted in Figure 1, the degradation may propagate within the spatial domain along certain directions. In this example, the propagation mainly follows the south-to-north direction, although a much weaker propagation along the west-east direction is also observed. This special phenomenon of spatial propagation, which is common in many engineering applications such as the surface erosion problems, needs to be addressed in the spatio-temporal degradation model, and is not considered by existing time-dependent stochastic degradation models.
- **anisotropy.** Because the propagation of degradation often presents a dominant direction,

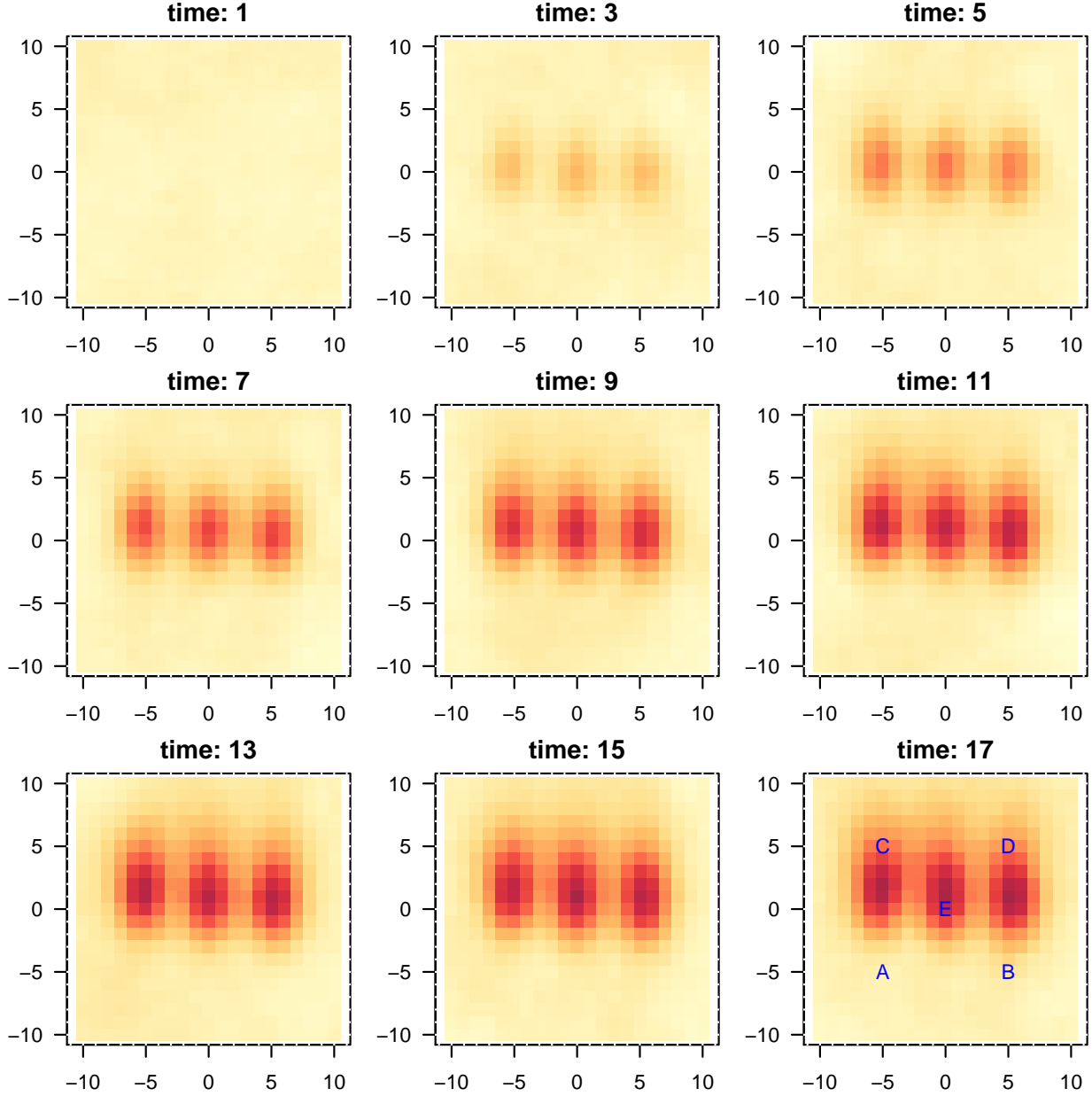


Figure 1: Measured degradation (the darker the color, the higher the degradation) on a two-dimensional surface at equally spaced time intervals. Three regions, aligned horizontally, are clearly observed with higher degradation level. The degradation propagates mainly along the south-to-north direction.

the measured degradation data over a spatial domain are directionally dependent, as opposed to isotropy which implies identical properties in all directions. From the statistical modeling point of view, if a random field is constructed to model the degradation data, the spatio-temporal correlation structure of that field is also expected to be anisotropic, making the

modeling much more challenging.

- **spatial heterogeneity.** Over a spatial domain, the degradation at different locations often exhibits a certain level of heterogeneity. The heterogeneity is due to not only the anisotropic degradation propagation discussed above, but also the difference in degradation rate at different spatial locations, governed by some underlying deterministic mechanism. In Figure 1, we clearly observe that the three high-degradation regions in the center have significantly higher degradation rate than other areas within the domain. Figure 2 shows the measured degradation, at different time points, for five fixed spatial locations (i.e., the observed marginal degradation at fixed locations). These five locations are marked by alphabets, “A”, “B”, “C”, “D” and “E”, in the bottom-right subplot of Figure 1. The heterogeneity is immediately seen. At locations “C”, “D” and “E”, the data show a clear degradation trend, and location “E” has the highest degradation rate among the three locations. At locations “A” and “B”, however, no significant degradation is observed, and the variation of the measured data at these two locations could be well attributed to random noise, such as measurement errors.

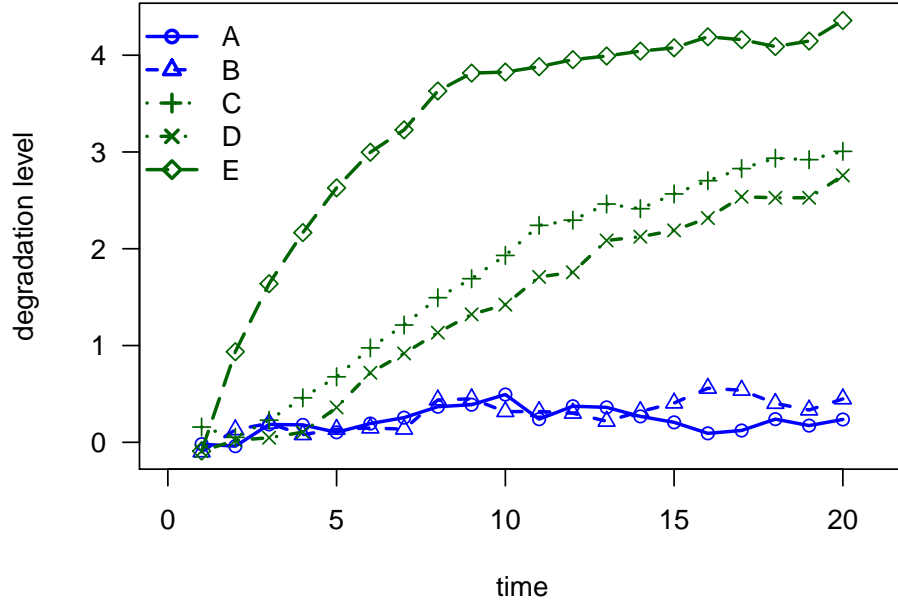


Figure 2: Measured degradation over time at five fixed spatial locations, marked by alphabets, “A”, “B”, “C”, “D” and “E”, in the bottom-right subplot of Figure 1.

- **spatio-temporal correlation.** The spatial propagation, anisotropy, and heterogeneity discussed above eventually lead to an extremely complex spatio-temporal correlation structure. For pure time-dependent stochastic degradation models, the temporal correlation structure can be directly specified, such as Brownian, Gamma and Inverse-Gaussian, and easily validated using data. For the modeling of spatio-temporal degradation data, however, one needs to choose not only the temporal correlation but also the spatial correlation structure. More importantly, the correlation is often anisotropic and space-time non-separable, meaning that

the spatio-temporal covariance function cannot be expressed as the multiplication of a spatial covariance function and a temporal covariance function. In fact, there have been prolonged interests in the spatio-temporal statistics to provide flexible and effective ways to construct covariance functions. Some key results are reported in Cressie and Huang (1999), Gneiting (2002), Banerjee et al. (2004), Fuentes et al. (2005) and Reich et al. (2011). For many real engineering applications, the space-time covariance structures can be extremely complicated due to the interactions between the spatial variation of degradation rate and degradation propagation over the entire spatial domain. It is challenging to specify appropriate space-time covariance functions that adequately model such complicated covariance structures (Calder, 2007; Ghosh et al., 2010).

- **computation cost.** The last challenge worth mentioning is the computation cost associated with parameter estimation. Because a spatio-temporal degradation dataset is usually large, estimating the unknown parameters of a spatio-temporal process can be computationally costly, if not impossible at all. For example, the Maximum Likelihood Estimation (MLE) typically requires an $\mathcal{O}((N_t \times N_s)^3)$ operation for the inversion of a large covariance matrix, where N_t and N_s are the total number of sampling times and the total number of sampling locations on the spatial domain. In addition, the total number of unknown parameters in a spatio-temporal degradation model is much larger than that of a pure time-dependent degradation model. Hence, maintaining the computation tractability for the proposed spatio-temporal degradation model is critical in order to make the proposed method practical.

1.3 Overview of the Paper

We provide an overview and describe the organization of the paper in this section. Section 2 presents the modeling framework as well as the key results of the spatio-temporal degradation model. We show that the proposed modeling approach leads to a random field with a space-time non-separable covariance structure. In particular, the degradation at a particular location and time is expressed as an additive superposition of two stochastic components: a dynamic spatial degradation generation process, and a spatio-temporal degradation propagation process. The first component, i.e., the spatial degradation generation process, is modeled as the sum a deterministic degradation generation process with (possibly time-varying) covariates and a white-in-time pure spatial process that accounts for the uncertainty associated with the degradation generation process. The second component, i.e., the spatio-temporal degradation propagation, is modeled through a convolution operation. Brown et al. (2000) advocated the use of convolution to approximate the propagation phenomenon in space under a constant vector field. Recently, Sigrist et al. (2015) presented a spatio-temporal Gaussian process directly derived from the solution of a stochastic PDE describing the convection-diffusion process, and proposed to use a Fourier spectral method for efficient computation. In spatial statistics, convolution was initially introduced as a novel way to construct the covariance function for complex non-stationary spatial processes (Higdon, 2002, 2007). In the same spirit, Calder (2007) proposed a dynamic space-time model in which the quantity of interest is expressed as the convolution of a latent process. Such a model fits into the general dynamic modeling framework for spatial-temporal data introduced in Stroud et al. (2001).

In Section 3, we show the connection between the proposed spatio-temporal degradation model and the existing pure time-dependent degradation model. In particular, when the spatial dependence of degradation is completely ignored, the proposed spatio-temporal degradation model

degenerates to a traditional Wiener degradation process at each spatial location. In Section 4 and Appendix B, we show the connection between the proposed spatio-temporal degradation model and physical PDE-based degradation models under special conditions. Such a connection not only helps to justify the proposed model, but also allows us to incorporate some physical knowledge into the statistical model for real engineering applications, if necessary. Recall that, one advantage of traditional time-dependent stochastic processes lies in its strong connection to stochastic PDE through which many physical degradation models are defined.

In Section 5, parameter estimation is firstly presented. Then, the motivating example in Section 1.2 is re-visited to demonstrate the proposed modeling approach. In particular, we provide discussions on how the model validation and selection can be performed based on data, as well as how the distributions of the First-Passage-Time and the First-Passage-Location can be approximated based on the simulation of the spatio-temporal degradation process.

Section 6 concludes the paper and highlights some future research directions.

2 A Spatio-Temporal Degradation Model

2.1 The Basic Framework

We first consider a discrete-in-time and continuous-in-space spatio-temporal random field, $\{Y(\mathbf{s}, t); \mathbf{s} \in \mathbb{R}^d, t \in \mathbb{N}^+\}$, where $Y(\mathbf{s}, t)$ represents the degradation at time t and location \mathbf{s} in a d -dimensional space. Without loss of generality, we let $d = 2$ which corresponds to the degradation over a two-dimensional surface, and the results presented in this paper can be directly extended to a higher dimension $d > 2$. As discussed in Section 1.1, the fundamental idea behind existing stochastic degradation models is rooted in the assumption of an additive accumulation of degradation. Following the same idea, we assume that $Y(\mathbf{s}, t)$ takes a general additive form as follows:

$$Y(\mathbf{s}, t) = G_\Delta(\mathbf{s}, t) + Z(\mathbf{s}, t), \quad (3)$$

where

$$G_\Delta(\mathbf{s}, t) = g_\Delta(\mathbf{s}, t) + \varepsilon_\Delta(\mathbf{s}, t), \quad (4)$$

with $Z(\mathbf{s}, t)$ being a stochastic process depending on $Y(\mathbf{s}, t - \Delta)$.

The first term $G_\Delta(\mathbf{s}, t)$ on the right hand side (RHS) of (3) is a spatial process that represents the amount of degradation *generated* at location \mathbf{s} over the time interval $(t - \Delta, t]$ with length Δ . Note that, one important difference between the spatio-temporal degradation and the well-studied pure time-dependent degradation is that the former problem requires us to consider, within the spatial domain, the source or the spatial variation of the degradation rate. Consider, for example, the surface degradation where the degradation is only initiated from some locations. To capture the spatial correlation of the generation of degradation within the spatial domain and account for the uncertainty over the time interval $(t - \Delta, t]$, $G_\Delta(\mathbf{s}, t)$ is further decomposed in (4) as the sum of a deterministic mean-value term $g_\Delta(\mathbf{s}, t)$ with (possibly time-varying) covariates, and a spatial process ε_Δ with covariance function $c_\Delta(\cdot) = \Delta \cdot c(\cdot)$, where $c(\cdot)$ is the covariance function of a white-in-time Gaussian process. We leave the detailed modeling of $g_\Delta(\mathbf{s}, t)$ to Section 2.2.

The second random term $Z(\mathbf{s}, t)$, which depends on $Y(\mathbf{s}, t - \Delta)$, captures a special but potentially important phenomenon in spatio-temporal degradation: *the propagation of degradation over space and time*. Some typical examples include the surface corrosion, crack propagation, etc., where

degradation is initiated at some locations and then propagated to neighboring areas along certain directions. It is possible to see that, the stochastic propagation process, $Z(\mathbf{s}, t)$, largely determines the spatio-temporal correlation structure of the random process $Y(\mathbf{s}, t)$. Following the work of Wikle and Cressie (1999), Brown et al. (2000) and Sigrist et al. (2015), we adopt the idea that the amount of degradation propagated to a certain location at time t can be expressed as a linear combination of the degradation at neighboring locations at time $t - \Delta$, weighted by some spatial Kernel function. Hence, a convolution model, with a Gaussian kernel, is used to describe the propagation process as follows (we provide discussions to justify the choice of the convolution approach and the Gaussian kernel in Sections 3 and 4):

$$\begin{aligned} Z(\mathbf{s}, t) &= \zeta_{\Delta} \{ \omega_{\Delta} * Y(\mathbf{s}, t - \Delta) \} \\ &= \zeta_{\Delta} \left\{ \int_{\mathbb{R}^2} \omega_{\Delta}(\mathbf{x}) Y(\mathbf{s} - \mathbf{x}, t - \Delta) d\mathbf{x} \right\} \\ &= \zeta_{\Delta} \{ \omega_{\Delta} * Z(\mathbf{s}, t - \Delta) + \omega_{\Delta} * g_{\Delta}(\mathbf{s}, t - \Delta) + \omega_{\Delta} * \varepsilon_{\Delta}(\mathbf{s}, t - \Delta) \}, \end{aligned} \quad (5)$$

where $*$ denotes the convolution operation, ω_{Δ} is the convolution kernel, and $\zeta_{\Delta} < 1$ is a scaling factor. In particular, we respectively define the scaling factor ζ_{Δ} and the convolution kernel ω_{Δ} as

$$\zeta_{\Delta} = \exp(-\lambda\Delta), \quad \lambda > 0, \quad (6)$$

$$\begin{aligned} \omega_{\Delta}(\mathbf{x}) &= \frac{1}{2\pi|\mathbf{\Sigma}_{\Delta}|^{1/2}} \exp \left\{ -\frac{(\mathbf{x} - \boldsymbol{\mu}_{\Delta})^{\top} \mathbf{\Sigma}_{\Delta}^{-1} (\mathbf{x} - \boldsymbol{\mu}_{\Delta})}{2} \right\} \\ &= \phi(\mathbf{x}; \boldsymbol{\mu}_{\Delta}, \mathbf{\Sigma}_{\Delta}), \end{aligned} \quad (7)$$

where ϕ denotes the probability density function of a bivariate Gaussian distribution, $\boldsymbol{\mu}_{\Delta}$ reflects the direction and speed of the spatial propagation of degradation over a time interval, and $\mathbf{\Sigma}_{\Delta}$ is the covariance matrix of the Gaussian convolution kernel.

The choice of the convolution model (5) as well as (6) and (7) is validated using data in the numerical example presented in Section 5, and is also theoretically rationalized in Sections 3 and 4 by establishing the connection between the proposed spatio-temporal model, respectively, and existing time-dependent degradation models and physical degradation models defined by stochastic PDE. Mathematically, the convolution model implies that the amount of degradation propagated to a location \mathbf{s} over a time interval is a linear combination of the degradation in the neighborhood of \mathbf{s} , given the direction and speed of the spatial propagation of degradation. Let \mathbf{v} be a vector that denotes the direction and speed of the spatial propagation of degradation over the spatial domain and over the time interval $(t - \Delta, t]$, we assume the following parameterization:

$$\boldsymbol{\mu}_{\Delta} = \mathbf{v}\Delta, \quad (8)$$

and

$$\mathbf{\Sigma}_{\Delta} = \mathbf{R}^{-1} \begin{pmatrix} \rho_1 \Delta & 0 \\ 0 & \rho_2 \Delta \end{pmatrix} (\mathbf{R}^{\top})^{-1}, \quad (9)$$

where \mathbf{R} is proper rotation matrix (i.e., $\mathbf{R}^\top = \mathbf{R}^{-1}$ and $\det \mathbf{R} = 1$) given by:

$$\mathbf{R} = \begin{pmatrix} \cos(\alpha_{\mathbf{v}}) & -\sin(\alpha_{\mathbf{v}}) \\ \sin(\alpha_{\mathbf{v}}) & \cos(\alpha_{\mathbf{v}}) \end{pmatrix} \quad (10)$$

with $\alpha_{\mathbf{v}} \in [0, 2\pi)$ being the counter-clockwise rotation angle of the propagation vector \mathbf{v} from the horizontal axis. Hence, the parameters, ρ_1 and ρ_2 in (9), respectively control the standard deviations of the convolution kernel ω_Δ in directions which are parallel and perpendicular to the direction of degradation propagation.

Further discretizing Δ in (5) into a number of n small time intervals ($\delta = \Delta/n$) gives

$$\begin{aligned} Z(\mathbf{s}, t) = & \sum_{i=1}^n \exp(-i\lambda\delta) \omega_\delta^{*i} * g_\delta(\mathbf{s}, t - i\delta) \\ & + \exp(-n\lambda\delta) \omega_\delta^{*n} * Z(\mathbf{s}, t - \Delta) + \sum_{i=1}^n \exp(-i\lambda\delta) \omega_\delta^{*i} * \varepsilon_\delta(\mathbf{s}, t - i\delta) \end{aligned} \quad (11)$$

with $*n$ denoting the n -fold convolution operation. Substituting (11) into (3) and noting that the convolution of Gaussians is still a Gaussian, we obtain the expression of the degradation over continuous space and discrete time:

$$\begin{aligned} Y(\mathbf{s}, t) = & \sum_{i=0}^n \{ \exp(-i\lambda\delta) \phi_i(\mathbf{s}) * g_\delta(\mathbf{s}, t - i\delta) \} \\ & + \sum_{i=0}^n \{ \exp(-i\lambda\delta) \phi_i(\mathbf{s}) * \varepsilon_\delta(\mathbf{s}, t - i\delta) \} \\ & + \exp(-n\lambda\delta) \phi_i(\mathbf{s}) * Z(\mathbf{s}, t - \Delta). \end{aligned} \quad (12)$$

where

$$\phi_i(\mathbf{s}) = \begin{cases} \phi(\mathbf{s}; i\boldsymbol{\mu}_\delta, i\boldsymbol{\Sigma}_\delta), & i \in \mathbb{N}^+ \\ \kappa(\mathbf{s}), & i = 0 \end{cases} \quad (13)$$

with κ being a Dirac delta function.

When $\lambda > 0$ and $n > 0$, $Y(\mathbf{s}, t)$ is approximated by a stationary spatio-temporal random field as follows:

$$\begin{aligned} Y(\mathbf{s}, t) \approx & \sum_{i=0}^{\infty} \{ \exp(-i\lambda\delta) \phi_i(\mathbf{s}) * g_\delta(\mathbf{s}, t - i\delta) \} \\ & + \sum_{i=0}^{\infty} \{ \exp(-i\lambda\delta) \phi_i(\mathbf{s}) * \varepsilon_\delta(\mathbf{s}, t - i\delta) \}. \end{aligned} \quad (14)$$

Let

$$\Psi_i(\mathbf{s}) = \exp(-i\lambda\delta) \phi_i(\mathbf{s}), \quad (15)$$

we have the following result that characterizes the spatio-temporal correlation structure of the

process defined in (14).

Lemma 1. *For the stochastic degradation process defined in (14), the covariance of the degradation between (\mathbf{s}_1, t_1) and (\mathbf{s}_2, t_2) (assuming $t_2 - t_1 = j\delta$ for some $j = 0, 1, 2, \dots$) is not space-time separable and is given by*

$$\begin{aligned} \text{cov}(Y(\mathbf{s}_1, t_1), Y(\mathbf{s}_2, t_2)) &= \sum_{i=0}^{\infty} (\tilde{\Psi}_i * \Psi_{j+i, t_2} * c_\delta)(\mathbf{d}) \\ &\quad + I_{\{j=0\}} c_\delta(\mathbf{d}) \end{aligned} \quad (16)$$

where $\mathbf{d} = \mathbf{s}_2 - \mathbf{s}_1$, $\tilde{\Psi}_i(\mathbf{s}) \equiv \Psi_i(-\mathbf{s})$, and $I_{\{j=0\}} = 1$ only when $j = 0$, otherwise, $I_{\{j=0\}} = 0$.

The derivation of (16) is provided in Appendix A. Note that, since the random field $\varepsilon(\mathbf{s}, t)$ is isotropic, $c(\mathbf{d}) = c(\|\mathbf{d}\|)$ with $\|\mathbf{d}\|$ being the distance between \mathbf{s}_1 and \mathbf{s}_2 . It is seen from (16) that the covariance, $\text{cov}(Y(\mathbf{s}_1, t_1), Y(\mathbf{s}_2, t_2))$, is determined by not only the separation of time and space, but also the degradation propagation speed and direction between times t_1 and t_2 .

2.2 A Linear Representation

As mentioned in the previous section, $g_\Delta(\mathbf{s}, t)$, with possibly time-varying covariates, is a deterministic mean-value term that represents the amount of degradation generated at location \mathbf{s} over the time interval $(t - \Delta, t]$. In fact, it is possible to show that the spatio-temporal degradation process $\{Y(\mathbf{s}, t)\}$ takes a linear form, if $g_\Delta(\mathbf{s}, t)$ is modeled as a linear function of its covariates.

Let

$$g_\Delta(\mathbf{s}, t) = \mathbf{x}_0(\mathbf{s}, t) \boldsymbol{\beta}^\top, \quad (17)$$

where $\mathbf{x}_0(\mathbf{s}, t) = (x^{(1)}(\mathbf{s}, t), \dots, x^{(k)}(\mathbf{s}, t))$ is a row vector of length k that contains covariates, and $\boldsymbol{\beta} = (b^{(1)}, \dots, b^{(k)})$ a row vector of parameters that determine the effects of covariates on the degradation generation over the time interval $(t - \Delta, t]$. Note that, the proposed model allows $\mathbf{x}_0(\mathbf{s}, t)$ to vary over time, and this is often the case when the covariates depend on dynamic environmental conditions.

Substituting the expression of $g_\Delta(\mathbf{s}, t)$ into (14), the expected degradation at location \mathbf{s} and time t is obtained:

$$\begin{aligned} \mathbb{E}(Y(\mathbf{s}, t)) &\approx g(\mathbf{s}, t) + \sum_{i=1}^n \left\{ \Psi_i(\mathbf{s}) * (\mathbf{x}_0(\mathbf{s}, t - i\delta) \boldsymbol{\beta}^\top) \right\} \\ &= \mathbf{x}_0(\mathbf{s}, t) \boldsymbol{\beta}^\top + \sum_{i=1}^n \left\{ \sum_{p=1}^k b^{(p)} \left[\Psi_i(\mathbf{s}) * x^{(p)}(\mathbf{s}, t - i\delta) \right] \right\}. \end{aligned} \quad (18)$$

Let $\mathbf{x}_i(\mathbf{s}, t) = (\Psi_i(\mathbf{s}) * x^{(1)}(\mathbf{s}, t - i\delta), \dots, \Psi_i(\mathbf{s}) * x^{(k)}(\mathbf{s}, t - i\delta))$ for $i = 0, \dots, n$, and recall that $\Psi_i(\mathbf{s}) = 1$ if $i = 0$, (18) is further simplified to a linear form:

$$\mathbb{E}(Y(\mathbf{s}, t)) = \left(\sum_{i=0}^n \mathbf{x}_i(\mathbf{s}, t) \right) \boldsymbol{\beta}^\top = \tilde{\mathbf{x}}(\mathbf{s}, t) \boldsymbol{\beta}^\top. \quad (19)$$

with $(\sum_{i=0}^n \mathbf{x}_i(\mathbf{s}, t)) = \tilde{\mathbf{x}}(\mathbf{s}, t)$.

Equation (19) implies that the covariates of $g_\Delta(\mathbf{s}, t)$ are firstly transformed by the convolution operation, and the expected degradation is given by a linear combination of the transformed covariates. Since convolution is essentially a linear operation, the linear form of (19) is a natural consequence when $g_\Delta(\mathbf{s}, t)$ is assumed to be linearly dependent on its covariates. Finally, suppose that degradation is measured at N_s locations and N_t sampling times, we obtain from (19) a linear form as follows:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad (20)$$

where $\mathbf{Y} = (Y(\mathbf{s}_1, t_1), \dots, Y(\mathbf{s}_{N_s}, t_1), \dots, Y(\mathbf{s}_{N_s}, t_{N_t}))^\top$ is a column vector of length $N_s \times N_t$ and $\mathbf{X} = (\tilde{\mathbf{x}}(\mathbf{s}_1, t_1), \tilde{\mathbf{x}}(\mathbf{s}_2, t_1), \dots, \tilde{\mathbf{x}}(\mathbf{s}_{N_s}, t_1), \dots, \tilde{\mathbf{x}}(\mathbf{s}_{N_s}, t_{N_t}))^\top$ is a $N_s N_t \times k$ matrix, and \mathbf{e} is the error process with mean zero and variance $\boldsymbol{\Sigma}_Y$ given by (16).

2.3 The Model under Continuous Space and Continuous Time Domain

To obtain the stochastic degradation process under continuous space and continuous time, $\{Y(\mathbf{s}, t); \mathbf{s} \in \mathbb{R}^d, t \in \mathbb{R}\}$, we extend the stochastic degradation process defined in (14). First, it is necessary to assume that there exists a real-value function $u(\mathbf{s}, t)$ that satisfies:

$$g_\delta(\mathbf{s}, t) = \int_{t-\delta}^t u(\mathbf{s}, x) dx. \quad (21)$$

Since $g_\delta(\mathbf{s}, t)$ is previously defined as the total amount of degradation generated at location \mathbf{s} and over the time interval $(t - \delta, t]$. Hence, $u_\delta(\mathbf{s}, t)$ can be naturally interpreted as the degradation generation rate at location \mathbf{s} and time t .

Similarly, we also assume that there exists a spatial random process such that the following stochastic integral holds:

$$\varepsilon_\delta(\mathbf{s}, t) = \int_{t-\delta}^t d\tau(\mathbf{s}, x) dx. \quad (22)$$

As shown in Brown et al. (2000), if $\delta^{-1}\varepsilon_\delta(\cdot, t)$ is asymptotically Gaussian with mean zero and covariance $c_\tau(\cdot)$, it is easy to see that $\tau(\mathbf{s}, t)$ in (22) is a spatially correlated Brownian motion such that $d\tau(\cdot, t) \sim N(0, c_\tau(\cdot)dt)$.

Hence, by letting $n \rightarrow \infty$ and $\delta \rightarrow 0$, we obtain from (14) the expression of the stochastic degradation process under continuous space and time:

$$\begin{aligned} Y(\mathbf{s}, t) &= (\Psi^{(c)} *_{\mathbf{s}} *_{t} u)(\mathbf{s}, t) + (\Psi^{(c)} *_{\mathbf{s}} *_{t} d\tau)(\mathbf{s}, t) \\ &= (\Psi^{(c)} *_{\mathbf{s}} *_{t} (u + d\tau))(\mathbf{s}, t) \end{aligned} \quad (23)$$

where $*_{\mathbf{s}}$ and $*_{t}$ are respectively the convolution with respect to space \mathbf{s} and time t , and $\Psi^{(c)}$, defined as follows, is the continuous version of $\Psi_i(\mathbf{s})$ in (15):

$$\Psi^{(c)}(\mathbf{s}, t) = \begin{cases} \exp(-\lambda t) \phi(\mathbf{s}; \boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t), & t \geq 0 \\ 0, & t < 0 \end{cases} \quad (24)$$

It is important to note that, $u + d\tau$ in (23) can be interpreted as the stochastic degradation rate which is expressed as the sum of a deterministic degradation rate u and a spatial error term $d\tau$. In fact, $u + d\tau$ is just the continuous version of equation (4), and equation (23) is the continuous version of (3) and (4). More importantly, (23) implies that the degradation at location \mathbf{s} and time t linearly depends, through the convolution operations in both space and time, on the degradation generated in the entire spatial domain and over the entire history $[0, t]$.

3 The Connection to Time-Dependent Degradation Models

In this section, we show the connection between the proposed spatio-temporal degradation model and the existing pure time-dependent stochastic degradation models. As discussed in the introduction section, existing degradation models can often be expressed by differential equations as follows:

$$dY(t) = udt + dB(t), \quad (25)$$

where $Y(t)$ is the degradation over time, u is the degradation rate, and $B(t)$ is some stochastic process; see Ye and Xie (2014) for a comprehensive review on stochastic degradation models under this framework. For example, if $B(t)$ is a Brownian process, (25) becomes the widely used Wiener degradation model.

Following the approach described in Brown et al. (2000), we now obtain the SPDEs representation of the spatio-temporal degradation model (23):

$$dY(\mathbf{s}, t) = (u(\mathbf{s}, t) - \frac{1}{2}\{\Gamma Y(\cdot, t)\}(\mathbf{s}))dt + d\tau(\mathbf{s}, t), \quad (26)$$

where Γ is a spatial linear operator defined as:

$$\{\Gamma f(\cdot)\}(\mathbf{s}) = \frac{\partial}{\partial \mathbf{s}^T} f(\mathbf{s})\boldsymbol{\mu} - \text{trace} \left\{ \frac{\partial^2}{\partial \mathbf{s} \partial \mathbf{s}^T} f(\mathbf{s}) \right\} \boldsymbol{\Sigma} + 2\lambda f(\mathbf{s}), \quad (27)$$

and $\tau(\mathbf{s}, t)$, previously defined in (22), is a spatially correlated Brownian motion such that $d\tau(\cdot, t) \sim N(0, c_\tau(\cdot)dt)$. Here, $c_\tau(\cdot) = \lim_{\delta \rightarrow 0}(\delta^{-1}c_\delta(\cdot))$.

It is possible to see that, the first term $u(\mathbf{s}, t)$ in (26) describes the degradation rate at location \mathbf{s} and at time t , which corresponds to the term u in (25). The second term $\frac{1}{2}\{\Gamma Y(\cdot, t)\}(\mathbf{s})$ captures the propagation of degradation over space, which of course does not appear in the traditional degradation model (25). The last stochastic term $d\tau(\mathbf{s}, t)$ in (26) describes the uncertainty associated with the degradation process, which naturally connects to $dB(t)$ in (25). Next, we consider two special scenarios:

- **The process without spatial degradation propagation and decay.** If the spatial degradation propagation and decay are ignored (i.e., $\boldsymbol{\mu} = \boldsymbol{\Sigma} = \lambda = 0$), the second term of (26) vanishes, and the degradation process (26) becomes

$$dY(\mathbf{s}, t) = u(\mathbf{s}, t)dt + d\tau(\mathbf{s}, t), \quad (28)$$

which is exactly a classic multivariate Wiener degradation process. For such a process, the correlation structure of this process is fully determined by the choice of $c(\mathbf{s})$. In the context

of spatial statistics, any general spatial covariance functions can be used such as Exponential, Matérn and Gaussian (Cressie and Huang, 1999). In particular, since $\tau(\mathbf{s}, t)$ is a spatially correlated Brownian motion (i.e., a multivariate Brownian motion), it is immediately seen that the (marginal) degradation process at any location \mathbf{s} becomes a traditional Brownian degradation process.

- **The process without spatial degradation propagation.** If only the spatial degradation propagation is ignored (i.e., $\boldsymbol{\mu} = \boldsymbol{\Sigma} = 0$ and $\lambda > 0$), the degradation process (26) becomes

$$dY(\mathbf{s}, t) = (u(\mathbf{s}, t) - \lambda Y(\mathbf{s}, t))dt + d\tau(\mathbf{s}, t). \quad (29)$$

From Sigrist et al. (2015), the process above is a space-time separable random process with covariance function $(2\lambda)^{-1} \exp(-\lambda|t|)c(\mathbf{s})$.

4 The Connection to Physical Models

For real engineering problems, physical degradation models are typically defined in the form of Partial Differential Equations (PDE); for example, the reaction-diffusion model for autocatalytic degradation in polymer microspheres (Versypt et al., 2015), the image quality degradation (Foyer and Zou, 2006), etc. Hence, it is of a great practical importance to establish the connection between the proposed statistical degradation model and physical degradation models given by PDE. Such a connection not only helps to justify the proposed statistical approach, such as the choice of the convolution approach with a Gaussian kernel (5), but also demonstrate how a physics-based statistical degradation model can be constructed for real problems.

In order to maintain the clarity and readability of the paper, the main results in this section are presented in Appendix B. In particular, we first establish the connection between the proposed model and a general PDE with convection, diffusion, decay and generation terms. Then, we illustrate such a connection using a real reaction-diffusion model that describes the gradual degradation (decomposition) of polymer microspheres composed of poly(D,L-lactic-co-glycolic acid) (PLGA). It can be seen that, the proposed statistical model is well justified when the direction and speed of degradation propagation are slowly varying within the spatial region.

5 Numerical Example

The motivating example presented in Section 1.2 is re-visited to demonstrate the modeling of spatio-temporal degradation data using the proposed approach. Four important issues are addressed, including the parameter estimation, discussions on the numerical results, model validation, and how the First-Passage-Time (FPT) and First-Passage-Location (FPL) can be obtained from numerical simulation.

5.1 Parameter Estimation

The proposed spatial-temporal degradation model contains a much larger number of parameters than the existing time-dependent degradation models. Here, the model parameters include: 1) the decay parameter λ , 2) the degradation propagation vector \mathbf{v} which contains both the speed and direction of propagation, 3) the parameters, ρ_1 and ρ_2 in (10), that control the standard deviations

of the convolution kernel, 4) the parameter $\boldsymbol{\theta}$ in the spatial covariance function $c(\cdot; \boldsymbol{\theta})$, and 5) the parameter $\boldsymbol{\beta}$, a row vector of length k in (17), that determines the effects of covariates on the degradation generation process. Note that, the parameters, λ , ρ_1 , ρ_2 and \mathbf{v} , determine the convolution kernel as well as the design matrix \mathbf{X} in the linear model (20). These parameters, together with $\boldsymbol{\theta}$, determine the covariance matrix of the error term \mathbf{e} in (20). And the last set of parameter $\boldsymbol{\beta}$ contains the coefficients of the linear model (20).

Although the number of parameters is large, the special structure of the proposed model allows us to estimate the unknown parameters all at once using the Maximum Likelihood Estimation (MLE). Following (3) and (4), we denote by

$$\tilde{\mathbf{Y}}(\mathbf{s}, t) = \mathbf{Y}(\mathbf{s}, t) - \mathbf{Z}(\mathbf{s}, t) = g_{\Delta}(\mathbf{s}, t) + \varepsilon_{\Delta}(\mathbf{s}, t) \quad (30)$$

the difference between the degradation at time t , and the degradation propagated to time t given the degradation at time $t - \Delta$. Since ε_{Δ} is a white-in-time Gaussian spatial process, $\tilde{\mathbf{Y}}(\mathbf{s}, t)$ defined above is also a white-in-time Gaussian spatial process with mean $g_{\Delta}(\mathbf{s}, t) = \mathbf{x}_0(\mathbf{s}, t)\boldsymbol{\beta}^{\top}$ and covariance matrix $\boldsymbol{\Sigma}_{\varepsilon}$. Here, the covariance matrix is determined through the covariance function $c_{\Delta}(\cdot) = \Delta \cdot c(\cdot)$, where $c(\cdot)$ is a covariance function of a white-in-time Gaussian process.

Let $\boldsymbol{\Omega}$ be the set that contains the model parameters, the log-likelihood function of $\boldsymbol{\Omega}$ given the observations $\tilde{\mathbf{Y}}(\mathbf{s}, t)$, for $t = 2, \dots, N_t$, are

$$\mathcal{L}(\boldsymbol{\Omega}) = \sum_{t=2}^{N_t} \left\{ \frac{1}{(2\pi)^{N_s/2} (\det \boldsymbol{\Sigma}_{\varepsilon})^{1/2}} \exp \left(- \frac{(\tilde{\mathbf{y}}(\mathbf{s}, t) - \mathbf{x}_0(\mathbf{s}, t)\boldsymbol{\beta}^{\top})^{\top} \boldsymbol{\Sigma}_{\varepsilon}^{-1} (\tilde{\mathbf{y}}(\mathbf{s}, t) - \mathbf{x}_0(\mathbf{s}, t)\boldsymbol{\beta}^{\top})}{2} \right) \right\}, \quad (31)$$

and the maximum likelihood estimator, $\hat{\boldsymbol{\Omega}}$, is found by maximizing the log-likelihood function (31).

5.1.1 Additional Remarks on Parameter Estimation

In general, estimating the unknown parameters of a spatio-temporal process can be computationally costly. Based on the linear model (17), \mathbf{Y} is a spatio-temporal process with its covariance structure fully characterized by (16). Hence, one might construct the likelihood function of the model parameter $\boldsymbol{\Omega}$ based on the observed \mathbf{y} , obtained from the entire spatial domain and the sampling time period. However, this is not at all practical for large data sets due to an $\mathcal{O}((N_t \times N_s)^3)$ cost for inverting the large covariance matrix (for example, $N_t = 20$ and $N_s = 441$ in the numerical example). Also note that, when N_s is large, the convolution operation itself on a two-dimensional surface can further slow down the numerical optimization.

Hence, we leverage the special structure given in (3) and (4), and obtain the likelihood function based on $\tilde{\mathbf{Y}}(\mathbf{s}, t)$ defined in (30), which is a white-in-time Gaussian spatial process. Following this approach, the computational cost is significantly reduced to $\mathcal{O}((N_t - 1) \times N_s^3)$, making the proposed model more practical. For the numerical example considered in this paper, the parameter estimation problem can be solved within an hour on a standalone machine. In fact, both the convolution operation on the two-dimensional space and the computation of the contribution to the total likelihood from the data collected at each time point can be *embarrassingly* parallelized with little effort needed to separate the problem into a number of parallel tasks.

In a special case when the values of λ , ρ_1 , ρ_2 and \mathbf{v} are known from some physical knowledge, the parameters, $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ associated with a linear model (17), can also be estimated using the Iteratively Re-Weighted Generalized Least Squares (IRWGLS) which consists of the following steps:

Step 1 : Set the initial $\hat{\Sigma}_Y$ to an identify matrix of size $N_s \times N_t$.

Step 2 : Estimate β using the Feasible General Least Squares (FGLS):

$$\hat{\beta} = (X^\top \hat{\Sigma}_Y^{-1} X)^{-1} X^\top \hat{\Sigma}_Y Y.$$

Step 3 : Based on $r = Y - X\hat{\beta} = (r(s_1, t_1), \dots, r(s_{N_s}, t_{N_t}))^\top$, estimate $\hat{\theta}$ of the parameter θ associated with the covariance function $c(\cdot; \theta)$, and obtain the estimate of the covariance matrix, $\hat{\Sigma}_Y$, from equation (16).

Step 4 : Iterate Steps 1 and 2 until the relative changes of $\hat{\beta}$ and $\hat{\theta}$ are small.

In the first iteration, since $\hat{\Sigma}_Y$ is an identify matrix, $\hat{\beta}$ in Step 1 is the Ordinary Least Squares (OLS) estimator and is unbiased. In subsequent iterations, the finite-sample properties of the FGLS estimator, $\hat{\beta}$, are usually unknown and can be studied case-by-case via Monte Carlo experiments. Asymptotically, the FGLS estimator possesses the asymptotic properties of the Maximum Likelihood estimator, and is equivalent to the Generalized Least Squares (GLS) estimator under regularity conditions (Schabenberger and Gotway, 2005). Note that, step 3 is computationally expensive and requires an $\mathcal{O}((N_t \times N_s)^3)$ operation if the MLE is used. Hence, one may adopt the fast cross-validation-type method described in Carroll et al. (1997). Specifically, let $r_{(-s_0, -t_0)}$ be a column vector of residuals with the residual at location s_0 and time t_0 removed, and let $\eta(s_0, t_0)$ be the leaving-one-station-out prediction error, i.e., error in predicting the residual $r(s_0, t_0)$ using only $r_{(-s_0, -t_0)}$ given by

$$\eta(s_0, t_0) = r(s_0, t_0) - \hat{r}(s_0, t_0),$$

where $\hat{r}(s_0, t_0)$, given below, is the well-defined Simple Kriging predictor which is known to be the best linear estimator of $r(s_0, t_0)$ under squared-error loss:

$$\hat{r}(s_0, t_0) = \gamma \Sigma_{r_{(-s_0, -t_0)}}^{-1} r_{(-s_0, -t_0)}.$$

Here, $\gamma = \text{cov}(r_{(-s_0, -t_0)}^\top, r(s_0, t_0))$, and the covariance matrix of $r_{(-s_0, -t_0)}$, $\Sigma_{r_{(-s_0, -t_0)}}$, can be calculated using equation (16).

The optimum value of θ is found by minimizing the sum of squared leaving-one-station-out prediction errors, i.e.,

$$\min_{\theta} \sum_i^{N_s} \sum_j^{N_t} \eta^2(s_i, t_j; \theta). \quad (32)$$

In summary, the MLE described in Section 5.1 is preferred as it is computationally efficient and the statistical properties of the ML estimator is well studied.

5.2 Numerical Results, Model Validation and Prediction

5.2.1 Estimated Model Parameters

We present the numerical results and validate the model in this section. Before the MLE is performed, one needs to firstly choose the spatial covariance function $c(\cdot)$ of the process ε in (4). In

this numerical example, three candidate covariance functions $c(\cdot)$, which are commonly used for stationary spatial Gaussian process, are assumed and the most appropriate covariance function is then selected based on the model validation. The three spatial covariance functions considered are as follows (Note that, the covariance function $c(\cdot)$ is not the covariance function of the spatio-temporal degradation process which is given in (16)):

- The Exponential covariance function:

$$c(d) = \theta_1 \exp(-d/\theta_2), \quad (33)$$

where d is the spatial distance, and θ_1 and θ_2 respectively explain the rate of decay and the scale of the spatial correlation.

- The Gaussian covariance function:

$$c(d) = \theta_1 \exp(-d^2/\theta_2), \quad (34)$$

where d is the spatial distance, and θ_1 and θ_2 respectively explain the rate of decay and the scale of the spatial correlation.

- The Matérn covariance function:

$$c(d) = \frac{\theta_1}{2^{\theta_3-1}\Gamma(\theta_3)} (2\theta_3^{1/2}d/\theta_2)^{\theta_3} \mathfrak{K}_{\theta_3}(2\theta_3^{1/2}d/\theta_2) \quad (35)$$

where \mathfrak{K}_{θ_3} is a modified Bessel function, the parameters θ_1 and θ_2 are respectively known as the sill and range parameters in spatial statistics, and θ_3 measures the degree of smoothness of the process ε . The Matérn covariance function incorporates the Exponential covariance function as its special case when $\theta_3 = 1/2$.

For this particular numerical example, it is known that the degradation generation within a time interval is proportional to some known covariates, hence, the length of the parameter β in (17) becomes one. Table 1 shows the estimated model parameters obtained from the MLE described in Section 5.1. It is seen from the table that all three models successfully capture the degradation propagation along the south-to-north direction, as the estimated first component of \mathbf{v} (i.e., the horizontal component) is extremely small. Note that, ρ_1 and ρ_2 in (9), respectively control the standard deviations of the convolution kernel ω_Δ in directions which are parallel and perpendicular to the direction of degradation propagation. We see, from Table 1, that the estimated value of ρ_1 is also greater than ρ_2 , indicating a higher level of uncertainty along the main degradation propagation direction. The estimated effect of degradation generation, β , is close to 1 from all three models. Before further analysis on the estimation results is possible, model validation is needed to select the most appropriate parametric form for the covariance function $c(\cdot)$.

5.2.2 Model Validation and Selection

To validate the model and select the best $c(\cdot)$, we again resort to the key observation that $\tilde{\mathbf{Y}}(\mathbf{s}, t)$ is a white-in-time Gaussian spatial process with mean $g_\Delta(\mathbf{s}, t) = \mathbf{x}_0(\mathbf{s}, t)\beta^\top$ and covariance matrix Σ_ε . Once the model parameters have been estimated, the residuals, $\mathbf{y}^*(\mathbf{s}, t) = \tilde{\mathbf{y}}(\mathbf{s}, t) - \mathbf{x}_0(\mathbf{s}, t)\hat{\beta}^\top$ for $t = 1, \dots, N_t$, should look like a number of N_t samples generated from a white-in-time Gaussian

Table 1: The estimated model parameters assuming different $c(\cdot)$

		covariance function $c(\cdot)$		
		Exponential	Gaussian	Matérn
parameters	λ	0.127	0.09	0.149
	\mathbf{v}	(-0.040,0.499)	(-0.004,0.793)	(0.006,0.598)
	ρ_1	1.119	2.247	0.802
	ρ_2	0.192	0.301	0.216
	θ_1	0.019	0.010	0.071
	θ_2	12.883	11.564	81.560
	θ_3	N/A	N/A	0.434
	β	0.977	1.251	1.108

spatial process ε_Δ with the covariance function $c_\Delta(\cdot) = \Delta \cdot c(\cdot)$. Hence, the model can be validated *graphically* by comparing the empirical semi-variogram estimated from the residuals $\mathbf{y}^*(\mathbf{s}, t)$ and the theoretical semi-variogram computed based on the estimated model parameters. This simple but elegant idea for model validation is similar to that of choosing the best probability distribution for lifetime data using probability plotting, which has been widely adopted in statistical reliability analysis (Meeker and Escobar, 1998). Also note that, this approach validates not only the choice of the covariance function $c(\cdot)$, but also the fitted degradation generation and propagation processes that yield the residuals.

To estimate the empirical semi-variogram of the residuals $\mathbf{y}^*(\mathbf{s}, t)$, the well-known Cressie-Hawkins robust estimator is used (Cressie and Hawkins, 1980). Since the residual, \mathbf{Y}^* , is a white-in-time process, we may write, for our problem, the Cressie-Hawkins estimator as follows:

$$\hat{\gamma}(d) = \left\{ \frac{1}{|N(d)|} \sum_{N(d)} \sum_{t=2}^{N_t} |\mathbf{y}^*(\mathbf{s}_i, t) - \mathbf{y}^*(\mathbf{s}_j, t)|^{1/2} \right\}^4 \times \left(0.914 + \frac{0.988}{|N(d)|} \right) \quad (36)$$

where $N(d) = \{(i, j) : |\mathbf{s}_i - \mathbf{s}_j| = d\}$ and $|N(d)|$ is the number of distinct elements in the set $N(d)$.

Figure 3 shows the comparison between the empirical semi-variogram and the theoretical semi-variogram computed based on the estimated model parameters, respectively assuming Exponential, Gaussian, and Matérn covariance functions. It is immediately seen that the best match between the empirical and the theoretical variogram is obtained if the Gaussian covariance function is chosen for $c(\cdot)$.

The general normality check of the residual process can be performed using the conventional chi-square q-q plot, and the plot is shown in Figure 4. Again, when the Gaussian covariance function is chosen for $c(\cdot)$, we see the best fit between the theoretical and sample quantiles of the residuals $\tilde{\mathbf{y}}(\mathbf{s}, t)$.

Finally, based on the results shown in Table 1 and the Gaussian covariance function chosen for $c(\cdot)$, we obtain the following observations:

- The amount of degradation propagated to its neighboring areas decays to 50% of its original value approximately after 8 time periods, given that $\hat{\lambda} = 0.09$.
- The spatial propagation of degradation follows the south-to-north direction, and the propagation speed is about 0.793 per unit time based on the estimated \mathbf{v} . In addition, a much

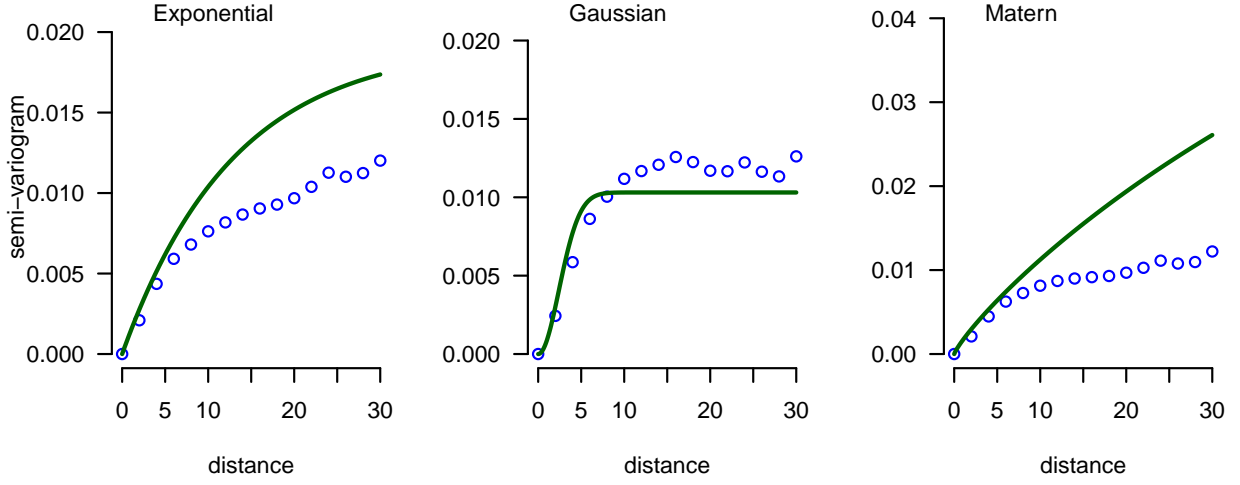


Figure 3: Comparison between the empirical semi-variogram estimated from the Cressie-Hawkins estimator and the theoretical semi-variogram computed based on the estimated model parameters, respectively assuming Exponential, Gaussian, and Matérn covariance functions for $c(\cdot)$.

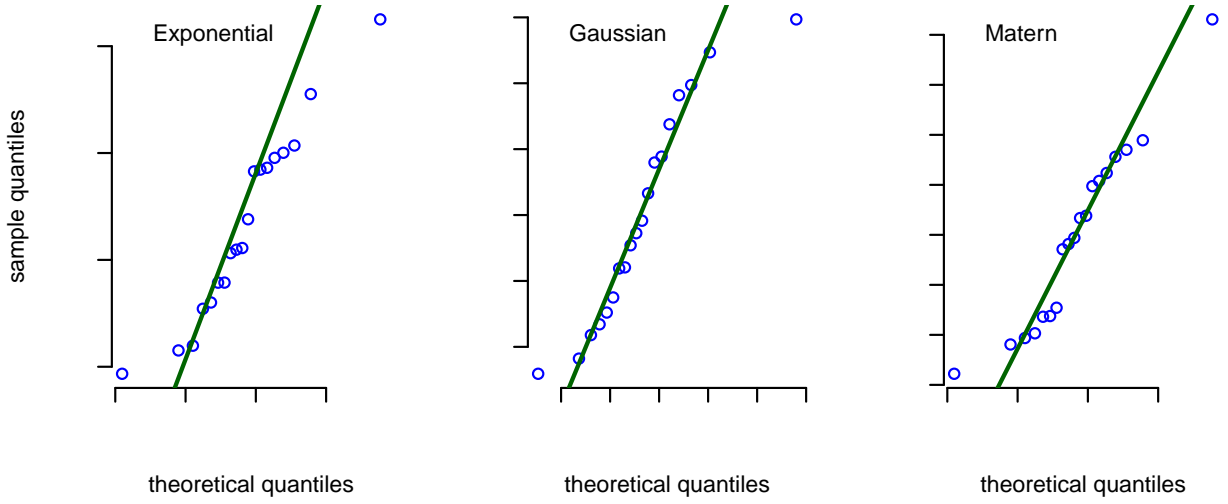


Figure 4: Chi-square q-q plot to check the multivariate normality of the residuals, respectively assuming Exponential, Gaussian, and Matérn covariance functions for $c(\cdot)$.

higher uncertainty is expected in the north-south direction than in the west-east direction, given that $\hat{\rho}_1$ is 7 times higher than $\hat{\rho}_2$.

- For the white-in-time Gaussian spatial process ε_Δ , the sill (i.e., the value at which the semi-variogram levels off) of its estimated semi-variogram is 0.01 based on $\hat{\theta}_1$, and the practical range (i.e., the spatial lag distance at which the semi-variogram reaches 95% of the sill) is

close to 6. Since the degradation data are aggregated to a 21×21 pixel array, the spatial process ε_Δ has a rather local effect on $G_\Delta(\mathbf{s}, t)$ in (4).

- The covariance function of the spatio-temporal degradation process $Y(\mathbf{s}, t)$ can now be calculated by (16). Figure 5 shows the computed covariance for different spatial and time lags. In particular, the two plots on the left respectively show the surface and contour plots of the covariance function for time lag 0, while the two plots on the right respectively show the surface and contour plots of the covariance function for time lag 2. It is seen that the covariance decreases as the time lag increases, as expected. For a fixed time lag, in particular, the covariance decreases faster in the horizontal direction as the degradation propagates vertically in this example.

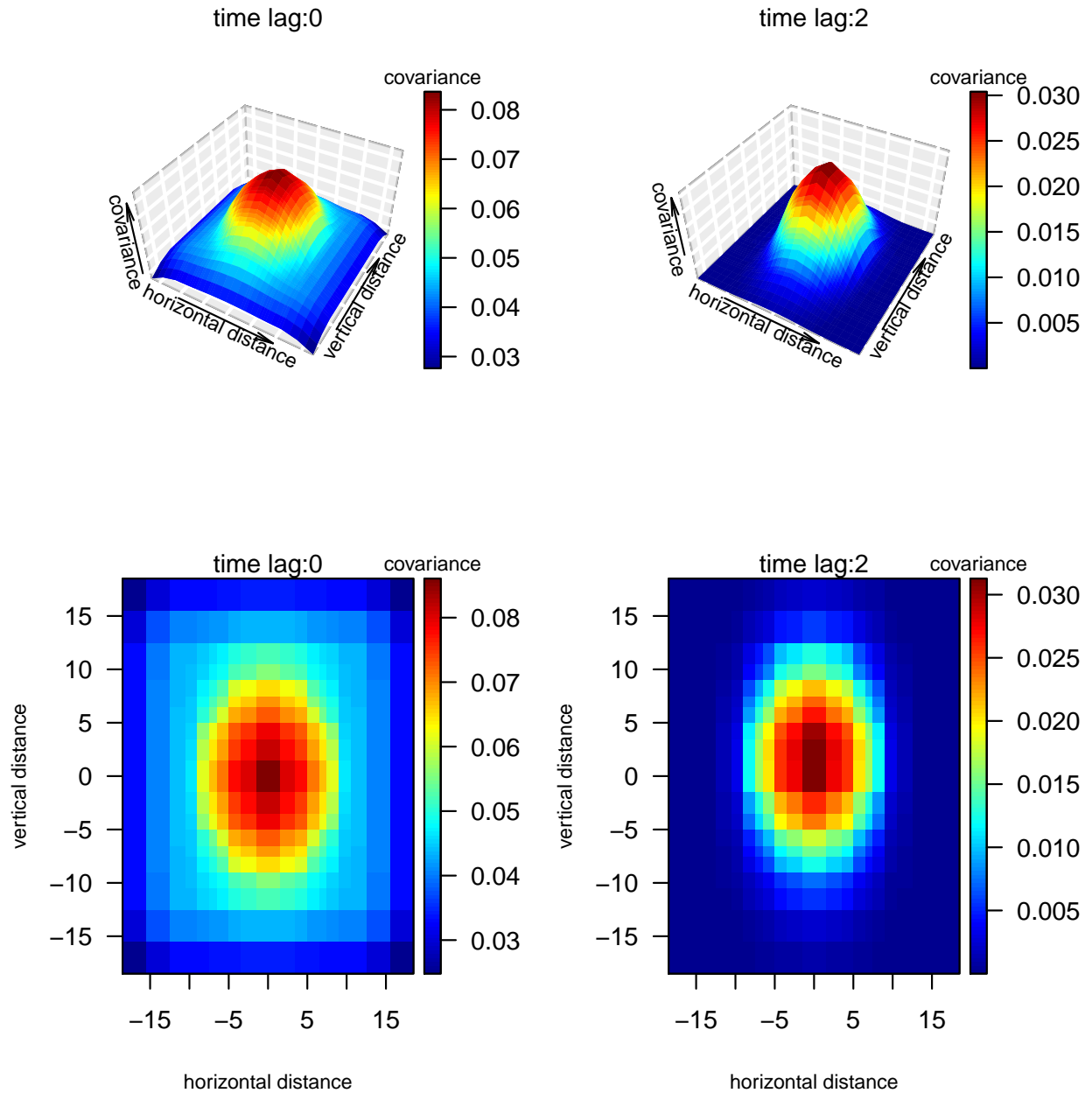


Figure 5: Covariance functions for different spatial and time lags. The two plots on the left respectively show the surface and contour plots of the covariance function for time lag 0, while the two plots on the right respectively show the surface and contour plots of the covariance function for time lag 2.

5.3 First-Passage-Time and First-Passage-Location

In degradation data analysis, the First-Passage-Time (FPT) of the degradation process for some given threshold is of great importance. The FPT is usually used to approximate the time-to-failure for reliability prediction, or, to determine the time for preventive maintenance. In the literature, the FPT for a pure time-dependent stochastic degradation process defined in (25) is given by

$$T^* = \inf\{t : Y(t) \geq y^*\}. \quad (37)$$

where y^* is the pre-specified threshold. In general, the closed-form expression of the distribution of T^* is not available. In some special cases when the degradation process is modeled by a Wiener process, it is well-known that the FPT follows the Inverse Gaussian distribution.

For a more complicated spatio-temporal degradation process, we are interested in not only when the degradation process hits the pre-determined threshold but also where it hits the threshold. In particular, the FPT for a spatio-temporal degradation process can be defined as:

$$T^* = \inf\{t : \max_{\mathbf{s}}(Y(\mathbf{s}, t)) \geq y^*\}. \quad (38)$$

Numerical simulation is needed to approximate the distribution of FPT. The simulation of random fields has been discussed in the literature (Lang and Potthoff, 2011; Brouste et al., 2007; Sigrist et al., 2015). For the proposed spatio-temporal model, the degradation process can be directly simulated using equations (3) and (4) in an iterative manner. This idea is similar to Sigrist et al. (2015) in which the authors investigate the generation of a Gaussian random field from stochastic PDE.

In particular, at any time t :

- compute the amount of degradation, $g_{\Delta}(\mathbf{s}, t)$, generated over the time interval $(t - \Delta, t]$ in equation (4).
- compute the propagation of degradation, $Z(\mathbf{s}, t)$, using the first line of equation (5). This involves a convolution operation on a two-dimensional space.
- simulate the noise term of equation (4), $\varepsilon_{\Delta}(\mathbf{s}, t)$.
- the simulated degradation at time t is computed by equations (3) and (4), i.e., the sum of the three terms computed in steps 1 to 3:

$$Y(\mathbf{s}, t) = g_{\Delta}(\mathbf{s}, t) + Z(\mathbf{s}, t) + \varepsilon_{\Delta}(\mathbf{s}, t).$$

For some given threshold value y^* , the left panel of Figure 6 shows the histogram and the density estimated from the kernel density estimation for the simulated Remaining Useful Life (RUL). The right panel of Figure 6 shows the reliability function estimated from the Kaplan-Meier method based on the simulated RUL. It is seen that there is a high probability that the degradation level will hit the threshold in the following 4 to 8 time periods. Next, Figure 7 shows the spatial distribution of the simulated First-Passage-Location (FPL). In this figure, the radius of each circle is proportional to the probability that the degradation will hit the threshold at the location where the circle resides. Since three regions with high degradation levels are observed in Figure 1, it is not surprising at all to note from Figure 7 that the FPL is most likely to be found in one of these three regions.

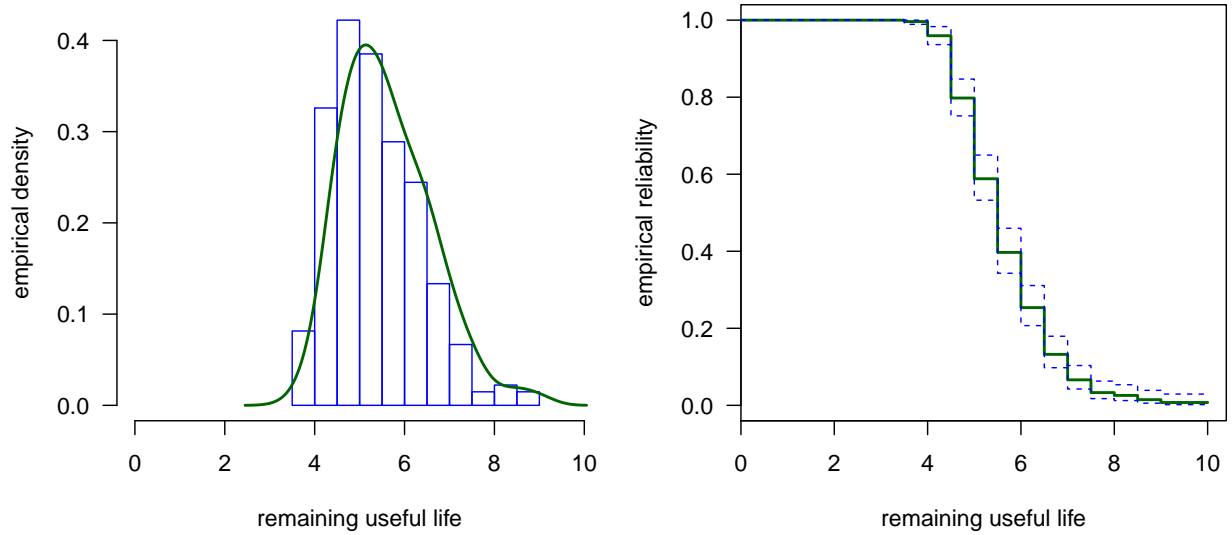


Figure 6: The histogram and the kernel density of the simulated Remaining Useful Life (on the left), and the reliability function estimated from the Kaplan-Meier method based on the simulated RUL (on the right).

6 Conclusions

This paper proposed a statistical degradation model for spatio-temporal degradation data. The proposed approach models the degradation process by a random field with a space-time non-separable and anisotropic covariance structure. The degradation at a particular location and time is expressed as the sum of a spatial degradation generation process, and a spatio-temporal degradation propagation process based on the convolution operation. Some unique challenges associated with the modeling of spatio-temporal degradation data were discussed, and a numerical example was presented to demonstrate the application of the proposed approach, including the parameter estimation, model validation and the approximations of the distribution of FPT and FPL using simulation. Note that, it has been assumed in this paper that the propagation of degradation is time-invariant, i.e., the degradation propagates along the same direction and at a constant speed. In some applications, it is possible that the propagation of degradation is affected by dynamic environmental conditions. In recent years, the modeling of pure time-dependent degradation data under dynamic environments has received much attention (Liao and Tian, 2012; Zhou, Serban and Gebraeel, 2014; Bian et al., 2015; Hong et al., 2015). Hence, extending the current modeling framework so as to incorporate the dynamic environments into the spatio-temporal degradation model is certainly an important topic and worth further investigating in the further research.

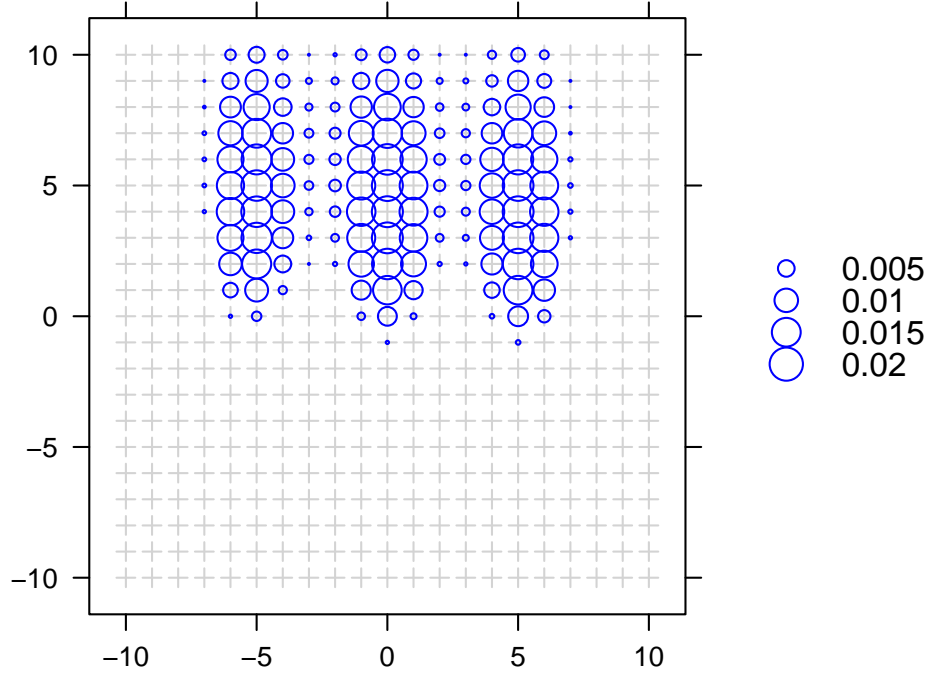


Figure 7: The spatial distribution of the simulated FPL. The radius of each circle is proportional to the probability (shown by the legend on the right) that the degradation will hit the threshold at the location where the circle resides.

7 Appendix A: Derivation of the Covariance (16)

In Appendix A, we show how equation (16) is derived. It follows from (14) that the covariance, $\text{cov}(Y(\mathbf{s}_1, t_1), Y(\mathbf{s}_2, t_2))$ for $t_1 \leq t_2$, can be written as:

$$\begin{aligned}
& \text{cov}(Y(\mathbf{s}_1, t_1), Y(\mathbf{s}_2, t_2)) = \\
& \text{cov}\left(\sum_{i=1}^{\infty} (\Psi_i(\mathbf{s}_1) * \varepsilon_{\delta}(\mathbf{s}_1, t_1 - i\delta)), \varepsilon_{\delta}(\mathbf{s}_2, t_2)\right) \\
& + \text{cov}(\varepsilon_{\delta}(\mathbf{s}_1, t_1), \varepsilon_{\delta}(\mathbf{s}_2, t_2)) \\
& + \text{cov}\left(\sum_{i=1}^{\infty} (\Psi_i(\mathbf{s}_2) * \varepsilon_{\delta}(\mathbf{s}_2, t_2 - i\delta)), \varepsilon_{\delta}(\mathbf{s}_1, t_1)\right) \\
& + \text{cov}\left(\sum_{i=1}^{\infty} (\Psi_i(\mathbf{s}_1) * \varepsilon_{\delta}(\mathbf{s}_1, t_1 - i\delta)), \right. \\
& \quad \left. \sum_{i=1}^{\infty} (\Psi_i(\mathbf{s}_2) * \varepsilon_{\delta}(\mathbf{s}_2, t_2 - i\delta))\right). \tag{39}
\end{aligned}$$

In what follows, we obtain the expression for each term on the right hand side (RHS) of (39). Since $\varepsilon_{\delta}(\mathbf{s}, t)$ is a white-in-time isotropic random field with spatial covariance function $c_{\delta}(\cdot)$, we immediately obtain the expressions of the first two terms on the RHS of (39):

$$\text{cov}\left(\sum_{i=1}^{\infty} (\Psi_i(\mathbf{s}_1) * \varepsilon_{\delta}(\mathbf{s}_1, t_1 - i\delta)), \varepsilon_{\delta}(\mathbf{s}_2, t_2)\right) = 0, \quad t_1 \leq t_2, \tag{40}$$

and

$$\text{cov}(\varepsilon_{\delta}(\mathbf{s}_1, t_1), \varepsilon_{\delta}(\mathbf{s}_2, t_2)) = \begin{cases} c_{\delta}(\mathbf{d}), & \text{if } t_1 = t_2 \\ 0, & \text{otherwise} \end{cases} \tag{41}$$

where $\mathbf{d} = \mathbf{s}_2 - \mathbf{s}_1$ is a vector. Note that, since the random field $\varepsilon_{\delta}(\mathbf{s}, t)$ is isotropic, $c_{\delta}(\mathbf{d}) = c_{\delta}(\|\mathbf{d}\|)$ with $\|\mathbf{d}\|$ representing the distance between \mathbf{s}_1 and \mathbf{s}_2 .

Let $t_1 = t_2 - j\delta$ for some $j \geq 0$, the expression of the third term on the RHS of (39) is derived as follows:

$$\begin{aligned}
& \text{cov}\left(\sum_{i=1}^{\infty} (\Psi_i(\mathbf{s}_2) * \varepsilon_{\delta}(\mathbf{s}_2, t_2 - i\delta)), \varepsilon_{\delta}(\mathbf{s}_1, t_1)\right) \\
& = \text{cov}(\Psi_{j,t_2}(\mathbf{s}_2) * \varepsilon_{\delta}(\mathbf{s}_2, t_1), \varepsilon_{\delta}(\mathbf{s}_1, t_1)) \\
& = \text{cov}\left(\int_{\mathbb{R}^2} \Psi_{j,t_2}(\mathbf{x}) \varepsilon_{\delta}(\mathbf{s}_2 - \mathbf{x}, t_1) d\mathbf{x}, \varepsilon_{\delta}(\mathbf{s}_1, t_1)\right) \\
& = (\Psi_{j,t_2} * c_{\delta})(\mathbf{d}). \tag{42}
\end{aligned}$$

The expression of the fourth term on the RHS of (39) can be derived in a similar way as follows:

$$\begin{aligned}
& \text{cov}\left(\sum_{i=1}^{\infty}(\Psi_i(\mathbf{s}_1) * \varepsilon_{\delta}(\mathbf{s}_1, t_1 - i\delta)), \sum_{i=1}^{\infty}(\Psi_i(\mathbf{s}_2) * \varepsilon_{\delta}(\mathbf{s}_2, t_2 - i\delta))\right) \\
&= \sum_{i=1}^{\infty} \text{cov}(\Psi_i(\mathbf{s}_1) * \varepsilon_{\delta}(\mathbf{s}_1, t_1 - i\delta), \Psi_{j+i, t_2}(\mathbf{s}_2) * \varepsilon_{\delta}(\mathbf{s}_2, t_2 - j\delta - i\delta)) \\
&= \sum_{i=1}^{\infty} (\tilde{\Psi}_i * \Psi_{j+i, t_2} * c_{\delta})(\mathbf{d})
\end{aligned} \tag{43}$$

where $\tilde{\Psi}_i(\mathbf{s}) \equiv \Psi_i(-\mathbf{s})$.

Note that, $\tilde{\Psi}_i$ and Ψ_i are Dirac delta functions when $i = 0$, we have

$$\begin{aligned}
\text{cov}(Y(\mathbf{s}_1, t_1), Y(\mathbf{s}_2, t_2)) &= (\Psi_{j, t_2} * c_{\delta})(\mathbf{d}) + \sum_{i=1}^{\infty} (\tilde{\Psi}_i * \Psi_{j+i, t_2} * c_{\delta})(\mathbf{d}) + I_{\{j=0\}} c_{\delta}(\mathbf{d}) \\
&= \sum_{i=0}^{\infty} (\tilde{\Psi}_i * \Psi_{j+i, t_2} * c_{\delta})(\mathbf{d}) + I_{\{j=0\}} c_{\delta}(\mathbf{d}).
\end{aligned} \tag{44}$$

where $I_{\{j=0\}} = 1$ only when $j = 0$, otherwise, $I_{\{j=0\}} = 0$.

8 Appendix B: The Connection to the Physical Degradation Model

The connection, between the proposed spatio-temporal degradation model and physical degradation models, is firstly established under a general setting. Then, we focus on a special case of a real physical degradation model.

Consider a general PDE which consists of four major components including generation, convection, diffusion and decay:

$$\frac{\partial \varphi(\mathbf{s}, t)}{\partial t} = Q(\mathbf{s}, t) - \nabla \cdot [\mathbf{v}(\mathbf{s}, t) \varphi(\mathbf{s}, t)] + \nabla \cdot [\mathbf{K}(\mathbf{s}, t) \cdot \nabla \varphi(\mathbf{s}, t)] - \frac{1}{\tau} \varphi(\mathbf{s}, t) \tag{45}$$

where φ is the quantity of interest at location \mathbf{s} and time t , Q is the generation rate, \mathbf{v} is the propagation velocity vector, \mathbf{K} is a second-order tensor of diffusivity, and τ is the relaxation timescale of decay. Note that, the second and the third terms on the RHS of (45) respectively represent the convection and diffusion processes.

First, we show that the convolution model (5) is an approximation to the physical convection-diffusion process under special conditions. Under a uniform and steady degradation propagation velocity field in a homogeneous space with zero degradation generation, (45) can be simplified as

$$\frac{\partial \varphi(\mathbf{s}, t)}{\partial t} = -\mathbf{v} \cdot \nabla \varphi(\mathbf{s}, t) + K \nabla^2 \varphi(\mathbf{s}, t) - \frac{1}{\tau} \varphi(\mathbf{s}, t). \tag{46}$$

Applying the Fourier transform, equation (46) becomes

$$\frac{d\tilde{\varphi}_{\eta}(t)}{dt} = -i(\boldsymbol{\eta} \cdot \mathbf{v}) \tilde{\varphi}_{\eta}(t) - \eta^2 K \tilde{\varphi}_{\eta}(t) - \frac{1}{\tau} \tilde{\varphi}_{\eta}(t), \tag{47}$$

in which η is the wave number and $\tilde{\varphi}_\eta$ is the Fourier coefficient of φ . It is not difficult to show that the solution of equation (47) is given by

$$\tilde{\varphi}_\eta(t + \Delta) = \exp \left\{ -\frac{1}{\tau} \Delta - [i(\boldsymbol{\eta} \cdot \mathbf{v}) + K\eta^2] \Delta \right\} \tilde{\varphi}_\eta(t), \quad (48)$$

and the backward Fourier transform leads to

$$\varphi(\mathbf{s}, t) = \exp \left\{ -\frac{1}{\tau} \Delta \right\} \int_{\mathbb{R}^2} \Omega_\Delta(\mathbf{x}) \varphi(\mathbf{s} - \mathbf{x}, t - \Delta) d\mathbf{x}, \quad (49)$$

in which the kernel $\Omega_\Delta(\mathbf{x})$ is exactly a Gaussian kernel:

$$\Omega_\Delta(\mathbf{x}) = \frac{1}{2\pi|\boldsymbol{\Sigma}_\Delta|^{1/2}} \exp \left\{ -\frac{(\mathbf{x} - \mathbf{v}\Delta)^\top \boldsymbol{\Sigma}_\Delta^{-1} (\mathbf{x} - \mathbf{v}\Delta)}{2} \right\}, \quad (50)$$

with the covariance matrix given by

$$\boldsymbol{\Sigma}_\Delta = \begin{pmatrix} 2K\Delta & 0 \\ 0 & 2K\Delta \end{pmatrix}. \quad (51)$$

The covariance matrix (51) is known as the diffusivity tensor in physics. The comparison between (6), (7) and (49) immediately justifies why Gaussian convolution kernel is chosen with $\lambda = \tau^{-1}$.

Equation (51) also suggests that

$$\rho_1 = 2K_\parallel(s, t), \quad \rho_2 = 2K_\perp(s, t) \quad (52)$$

where K_\parallel and K_\perp are the diffusivities respectively parallel and perpendicular to the propagation direction. In a special case when an isotropic diffusion model is used (i.e., $K = K_\parallel = K_\perp$), the covariance matrix reduces to equation (51). The connection shown above shows that the proposed statistical model is well motivated when *the direction and speed of degradation propagation are slowly varying within a certain spatial region*.

To see the relationship between $g_\Delta(\mathbf{s}, t)$ and the source term $Q(\mathbf{s}, t)$ in (45), we now consider the limit situation with zero diffusion and no decay, i.e., $\mathbf{K} \rightarrow 0$ and $\tau \rightarrow \infty$. Then, equation (45) reduces to

$$\frac{\partial \varphi(\mathbf{s}, t)}{\partial t} = -\nabla \cdot [\mathbf{v}(\mathbf{s}, t) \varphi(\mathbf{s}, t)] + Q(\mathbf{s}, t). \quad (53)$$

It is obvious that $\zeta_\Delta = 1$ and the convolution kernel ω_Δ in (6) in this limit becomes a Dirac delta function,

$$\omega_\Delta(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{v}\Delta).$$

Hence, substituting (5) into (3) and (4), and omitting the error process yields

$$Y(\mathbf{s}, t) = \{\omega_\Delta * Y(\mathbf{s}, t - \Delta)\} + g_\Delta(\mathbf{s}, t) = Y(\mathbf{s} - \mathbf{v}\Delta, t - \Delta) + g_\Delta(\mathbf{s}, t). \quad (54)$$

Since $Y(\cdot, t)$ in the statistical model corresponds to the quantity φ at t in the physical model, equation (54) can be rewritten as

$$\varphi(\mathbf{s}, t) = \varphi(\mathbf{s} - \mathbf{v}\Delta, t - \Delta) + g_\Delta(\mathbf{s}, t). \quad (55)$$

Applying Taylor expansion and assuming the mass conservation (i.e., $\nabla \cdot \mathbf{v}(\mathbf{s}, t) = 0$), equation (55) yields

$$\Delta \left(\frac{\partial \varphi(\mathbf{s}, t)}{\partial t} + \nabla \cdot [\mathbf{v}(\mathbf{s}, t) \varphi(\mathbf{s}, t)] - [g_\Delta(\mathbf{s}, t)/\Delta] \right) = O(\Delta^2). \quad (56)$$

By comparing (53) and (56), it is clear that the convolution model (54) is a first-order approximation to the scalar transport equation (53), i.e., (56) converges to (53) as $\Delta \rightarrow 0$. Hence, the amount of degradation generated over a time interval $(t - \Delta, t]$, $g_\Delta(\mathbf{s}, t)$, in the convolution model (5), directly corresponds to the generation term in the physics model (45) through the following relationship: $g_\Delta(\mathbf{s}, t) = Q(\mathbf{s}, t)\Delta$. And $u(\mathbf{s}, t)$ introduced in (21) becomes exactly the degradation rate.

Based on the results above, we now consider a recently proposed reaction-diffusion model that describes the gradual degradation (decomposition) of polymer microspheres composed of poly(D,L-lactic-co-glycolic acid) (PLGA) (Versypt et al., 2015). The PLGA microspheres are used for pharmaceutical drug delivery over extended periods of time. Compared to conventional drug dosage forms, the PLGA microspheres are biodegradable polymeric devices for controlled-release drug delivery, which provide enhanced control of drug concentrations and biodistribution, reduce side effects, and improve patient compliance. The reaction-diffusion model for the degradation of polymer microspheres within a radially symmetric sphere is given by

$$\frac{\partial \varphi(\mathbf{r}, t)}{\partial t} = Q(\mathbf{r}, t) + K \nabla^2 \varphi(\mathbf{r}, t) \quad (57)$$

where $\varphi(\mathbf{r}, t) = rc(\mathbf{r}, t)$ with $0 \leq \mathbf{r} \leq 1$ and $c(\mathbf{s}, t)$ respectively being the normalized radial position and concentration, K is the normalized diffusion coefficient, and $Q(\mathbf{r}, t)$ is the net reaction rate of generation of species per volume (interested readers may refer to Versypt et al. (2015) for more details). Note that, (57) is a typical degradation model with a generation and a diffusion term.

Consider a pure diffusion process with $Q(\mathbf{r}, t) = 0$, (46) becomes

$$\frac{\partial \varphi(\mathbf{r}, t)}{\partial t} = K \nabla^2 \varphi(\mathbf{r}, t). \quad (58)$$

Applying the Fourier transform, equation (58) becomes

$$\frac{d\tilde{\varphi}_\eta(t)}{dt} = -\eta^2 K \tilde{\varphi}_\eta(t), \quad (59)$$

in which η is the wave number and $\tilde{\varphi}_\eta$ is the Fourier coefficient of φ . And from (49), we have

$$\varphi(\mathbf{s}, t) = \int_{\mathbb{R}^2} \Omega_\Delta(\mathbf{x}) \varphi(\mathbf{r} - \mathbf{x}, t - \Delta) d\mathbf{x}, \quad (60)$$

in which the kernel $\Omega_\Delta(\mathbf{x})$ is exactly a Gaussian kernel:

$$\Omega_\Delta(\mathbf{x}) = \frac{1}{2\pi|\Sigma_\Delta|^{1/2}} \exp \left\{ -\frac{\mathbf{x}^\top \Sigma_\Delta^{-1} \mathbf{x}}{2} \right\}, \quad (61)$$

with the covariance matrix given by

$$\mathbf{\Sigma}_\Delta = \begin{pmatrix} 2K\Delta & 0 \\ 0 & 2K\Delta \end{pmatrix}. \quad (62)$$

Based on (60), the Gaussian convolution kernel can be chosen for the statistical model as shown in (6) and (7). Equation (62) also suggests that $\rho_1 = \rho_2 = 2K(\mathbf{r}, t)$.

Finally, it follows from (56) that

$$\Delta \left(\frac{\partial \varphi(\mathbf{s}, t)}{\partial t} - [g_\Delta(\mathbf{s}, t)/\Delta] \right) = O(\Delta^2). \quad (63)$$

Hence, as $\Delta \rightarrow 0$, $g_\Delta(\mathbf{s}, t)$ in the statistical degradation model is exactly the reaction term in the physical model (57) proposed by Versypt et al. (2015).

References

- Bae, S. J. and Kvam, P. H. (2004), A Nonlinear Random-Coefficients Model for Degradation Testing. *Technometrics*, **46**, 460–469.
- Banerjee, S., Carlin, B.P. and Gelfand, A.E. (2004), Hierarchical Modeling and Analysis for Spatial Data, Chapman & Hall/CRC, New York.
- Bhattacharyya, G. K. and Fries, A. (1982), Fatigue Failure Models-Birnbaum-Saunders vs. Inverse Gaussian. *IEEE Transactions on Reliability*, **31**, 439–440.
- Bian, L. K., Gebraeel, N. and Kharoufeh, J. P. (2015), Degradation Modeling for Real-Time Estimation of Residual Lifetimes in Dynamic Environments. *IIE Transaction*, **47**, 471CC486.
- Brouste, A., Istas, J. and Lambert-Lacroix, S. (2007), On Fractional Gaussian Random Fields Simulation. *Journal of Statistical Software*, **23**.
- Brown, P.E., Karesen, K.F., Roberts, G.O. and Tonellato, S. (2000), Blur-Generated Non-Separable Space-Time Models. *Journal of the Royal Statistical Society, Ser. B*, **62**, 847–860.
- Calder, C. (2007), Dynamic Factor Process Convolution Models for Multivariate Space-Time Data with Application to Air Quality Assessment. *Environmental and Ecological Statistics*, **14**, 229–247.
- Carroll, R., Chen, E., Li, T., Newton, H., Schmiediche, H. and Wang, N. (1997), Ozone Exposure and Population Density in Harris County, Texas. *Journal of the American Statistical Association*, **92**, 392–404.
- Chen, N. and Tsui, K. L. (2012), Condition Monitoring and Remaining Useful Life Prediction Using Degradation Signals: Revisited. *IIE Transaction*, **45**, 939–952.
- Cressie, N. and Hawkins, D. (1980), Robust Estimation of the Variogram. *Journal of the International Association for Mathematical Geology*, **12**, 115–125.
- Cressie, N. and Huang, H. (1999), Classes of Nonseparable Spatiotemporal Stationary Covariance Functions. *Journal of the American Statistical Association*, **94**, 1330–1340.
- Doganaksoy, N. and Hall, D. (2013), Gaining Physical Insights from Degradation Data. *Journal of Quality Technology*, **45**, 188–199.
- Doksum, K. A. and Hoyland, A., (1992), Models for Variable-Stress Accelerated Life Testing Experiments Based on Wiener Processes and the Inverse Gaussian Distribution. *Technometrics*, **34**, 74–82.
- Foyer, M. and Zou, D., (2006), A Global Approach for Solving Evolutive Heat Transfer for Image Denoising and Inpainting. *IEEE Transactions on Image Processing*, **15**, 2558–2573.
- Fuentes, M., Chen, L., Davis, J. and Lackmann, G. (2005), A New Class of Nonseparable and Nonstationary Covariance Models for Wind Fields. *Environmetrics*, **16**, 449–464.
- Ghosh, S., Bhave, P., Davis, J. and Lee, H. (2010), Spatio-Temporal Analysis of Total Nitrate Concentrations Using Dynamic Statistical Models. *Journal of the American Statistical Association*, **105**, 538–551.
- Gneiting, T. (2002), Nonseparable Stationary Covariance Functions for Space-Time Data. *Journal of the American Statistical Association*, **97**, 590–600.
- Higdon, D. (2002), Space and Space-Time Modeling Using Process Convolutions. in *Quantitative Methods for Current Environmental Issues*, eds. Anderson, C., Barnett, V., Chatwind, P., and El-Shaarawi, A., Springer Verlag, 37–56.

- Higdon, D. (2007), A Process-Convolution Approach to Modeling Temperatures in the North Atlantic Ocean. *Environmental and Ecological Statistics*, **5**, 173–190.
- Hong, Y., Duan, Y., Meeker, W. Q., Stanley, D. L. and Gu, X. (2015), Statistical Methods for Degradation Data with Dynamic Covariates Information and an Application to Outdoor Weathering Data. *Technometrics*, **57**, 180–193.
- Lawless, J. F. and Crowder, M. J. (2004), Covariates and Random Effects in a Gamma Process Model with Application to Degradation and Failure. *Lifetime Data Analysis*, **10**, 213–227.
- Liao, H. T. and Tian, Z. G. (2012), A Framework for Predicting the Remaining Useful Life of a Single Unit under Time-Varying Operating Conditions. *IIE Transactions*, **45**, 964–980.
- Lang, A. and Potthoff, J. (2011), Fast simulation of Gaussian random fields. *Monte Carlo Methods and Applications*, **17**, 195–214.
- Li, M. and Meeker, W. (2013), Application of Bayesian Methods in Reliability Data Analysis. *Journal of Quality Technology*, **46**, 1–23.
- Liu, X. and Tang, L. C. (2013), A Bayesian Optimal Design for Accelerated Degradation Tests. *Quality and Reliability Engineering International*, **26**, 863–875.
- Lu, C. J., Meeker, W. Q. and Escobar, L. A. (1996), A Comparison of Degradation and Failure-Time Methods for Estimating a Time-to-Failure Distribution. *Statistica Sinica*, **6**, 531–546.
- Meeker, W. Q. and Escobar, L. A. (1998), *Statistical Methods for Reliability Data*, John Wiley & Sons, New York.
- Meeker, W. Q., Escobar, L. A. and Lu, C. J. (1988), Accelerated Degradation Tests: Modeling and Analysis. *Technometrics*, **40**, 89–99.
- Nikulin, M. S., Limnios, N., Balakrishnan, N., Kahle, W. and Huber-Carol, C. (2010), *Advances in Degradation Modeling: Applications to Reliability, Survival Analysis, and Finance*, Springer Science & Business Media, New York.
- Peng, C. Y. and Tseng, S. T. (2010), Progressive-Stress Accelerated Degradation Test for Highly-Reliable Products. *IEEE Transactions on Reliability*, **59**, 30–37.
- Reich, B., Eidsvik, J., Guindani, M., Nail, A. and Schmidt, A. (2011), A Class of Covariate-Dependent Spatiotemporal Covariance Functions for the Analysis of Daily Ozone Concentration. *Annals of Applied Statistics*, **5**, 2425–2447.
- Schabenberger, O. and Gotway, C. (2005), *Statistical Methods for Spatial Data Analysis*, Chapman & Hall/CRC, Boca Raton.
- Sigrist, F., Kunsch, H. R. and Stahel, W. A. (2015). Stochastic Partial Differential Equation based Modelling fo Large Space-Time Data Sets. *Journal of the Royal Statistical Society, Ser. B*, **77**, 3–33.
- Sigrist, F., Kunsch, H. and Stahel, W. (2015), spate: An R Package for Spatio-Temporal Modeling with a Stochastic Advection-Diffusion Process. *Journal of Statistical Software*, **63**.
- Singpurwalla, N. (1995), Survival in Dynamic Environments. *Statistical Science*, **10**, 86–103.
- Stroud, J., Muller, P. and Sanso, B. (2001), Dynamic Models for Spatiotemporal Data. *Journal of the Royal Statistical Society, Ser.B*, **63**, 673–689.
- Tseng, S. T. and Peng, C. Y. (2004), Optimal Burn-In Policy by Using an Integrated Wiener Process. *IIE Transactions*, **36**, 1161–1170.

Versypt, A. N. F., Arendt, P. D., Pack, D. W. and Braatz, R. D. (2015), Derivation of an Analytical Solution to a Reaction-Diffusion Model for Autocatalytic Degradation and Erosion in Polymer Microspheres. *PLoS ONE*, **10**, e0135506, doi:10.1371.

Wikle, C. K. and Cressie, N. (1999), A Dimension-Reduced Approach to Space-Time Kalman Filtering. *Biometrika*, **86**, 815–829.

Ye, Z. S. and Xie, M. (2014), Stochastic Modelling and Analysis of Degradation for Highly Reliable Products. *Applied Stochastic Models in Business and Industry*, **31**, 16–32.

Ye, Z. S. and Chen, N. (2014), The Inverse Gaussian Process as a Degradation Model. *Technometrics*, **56**, 302–311.

Zhou, R. S., Serban, N. and Gebraeel, N. (2014), Degradation-Based Residual Life Prediction under Different Environments. *The Annals of Applied Statistics*, **8**, 1671–1689.